

# 美团即时智能配送平台高可用实践

侯现佑 201605



# 目录

.....

## 即时智能配送

---

- 业务特点
- 美团配送发展历程
- 业务概览

## 高可用架构

---

- 高可用架构
- 双机房容灾

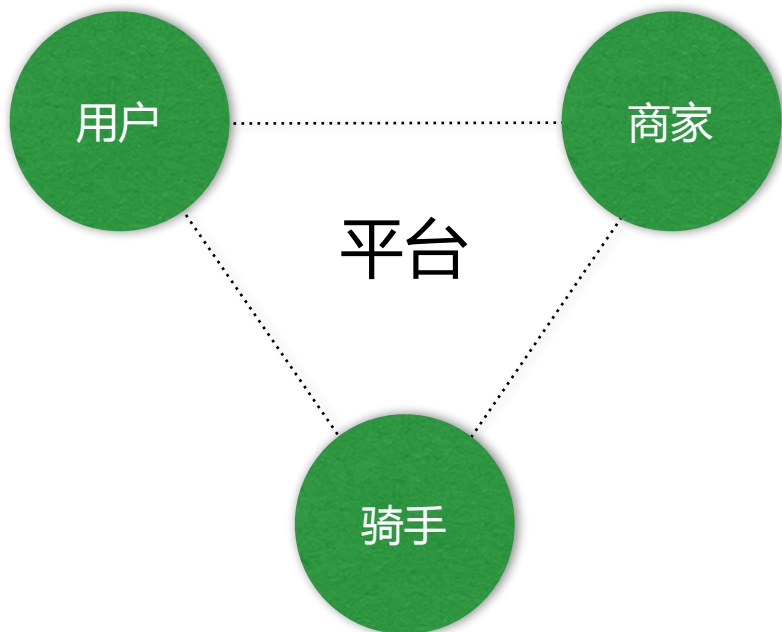
## 系统高可用实践

---

- 容量规划
- 降级
- 限流
- 监控报警体系
- 流程保障

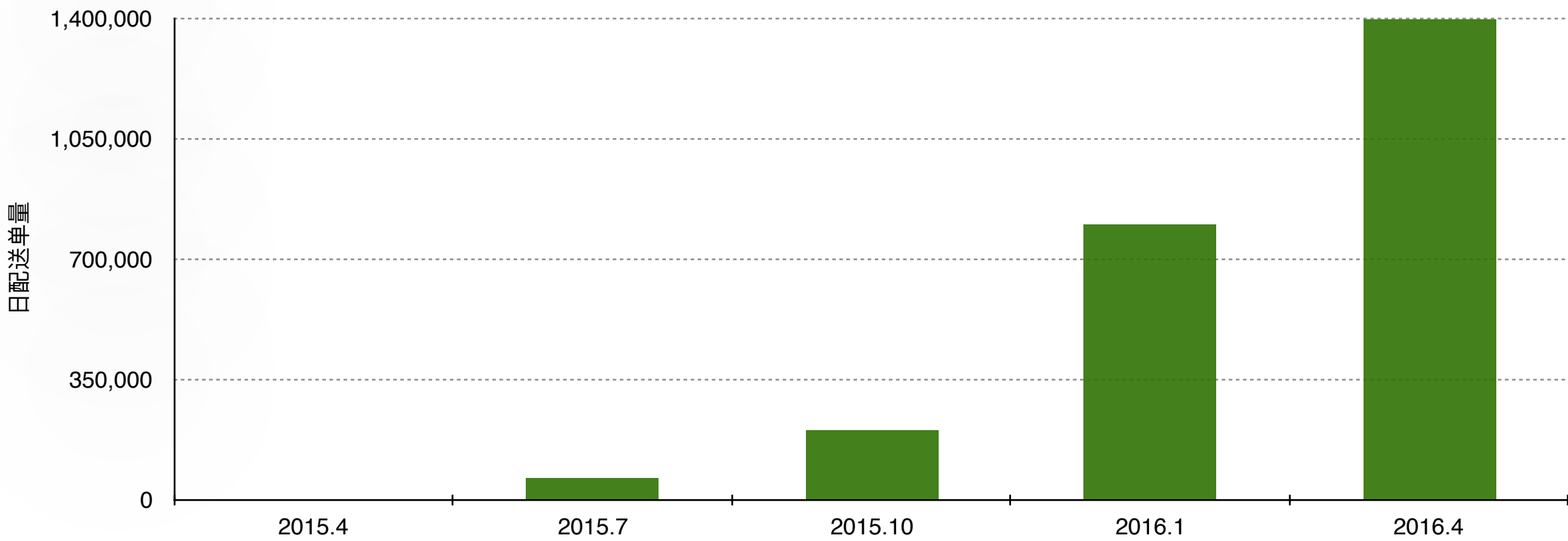
# 即时智能配送

---



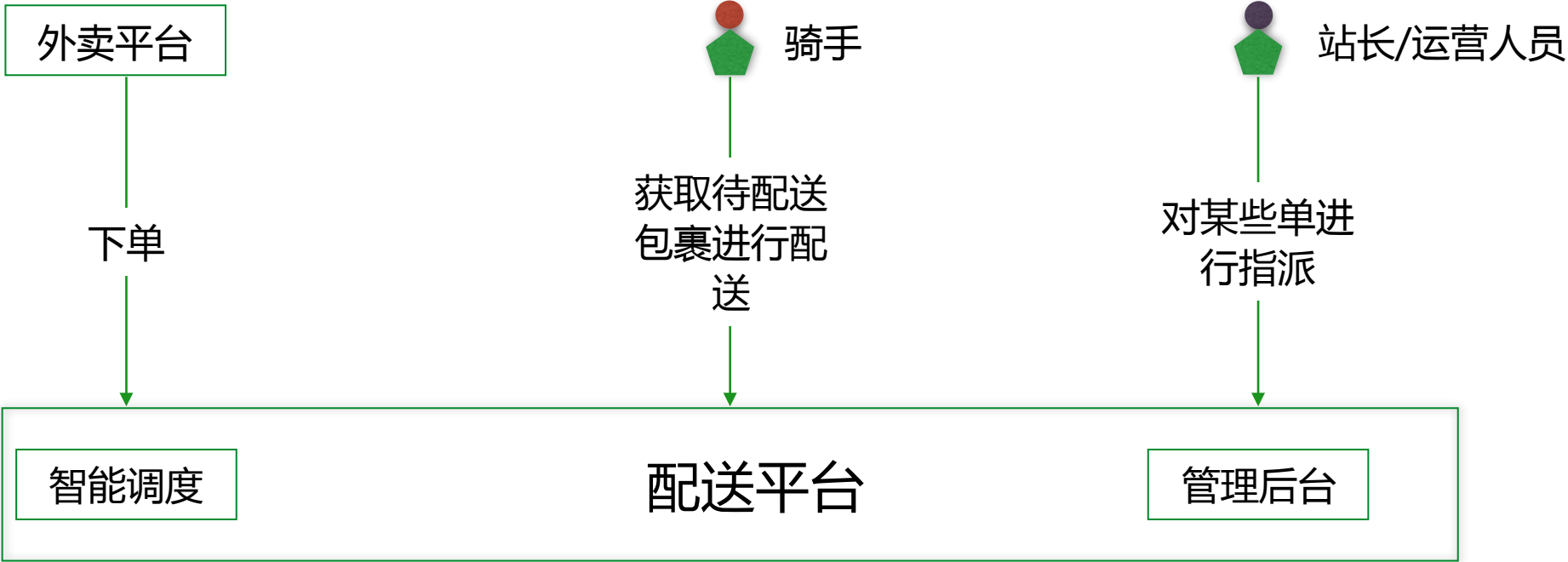
- 业务复杂：四方利益、三条业务线
- 可用性要求高
- 业务形态存在明显尖峰
- 对及时性要求高
- 运力受限于线下
- 算法要求高：多目标运筹优化

# 美团配送发展

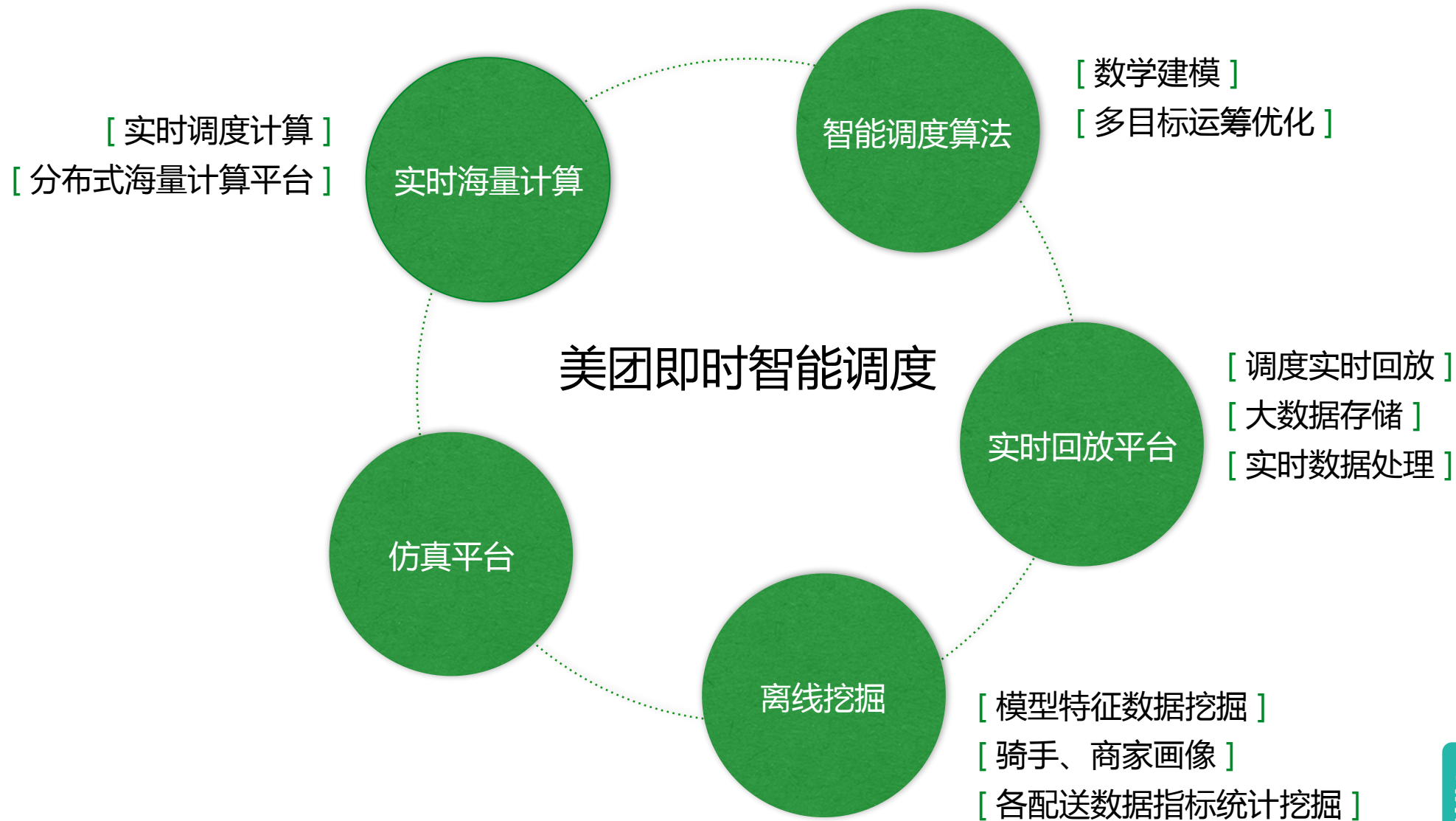


# 美团配送业务1：主交易流程

.....



# 美团配送业务2：智能调度

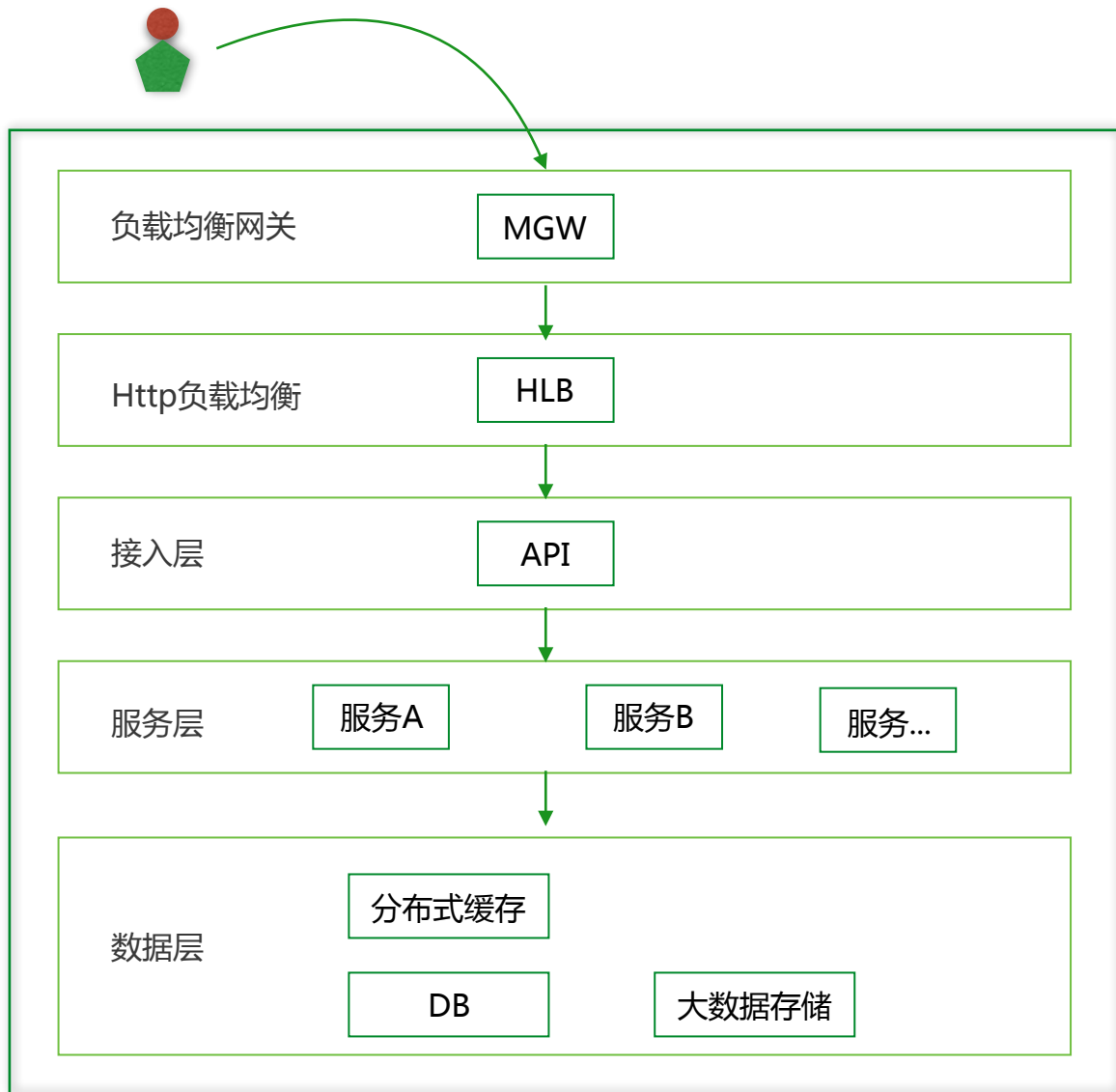


高可用：架构

高可用架构

双机房容灾

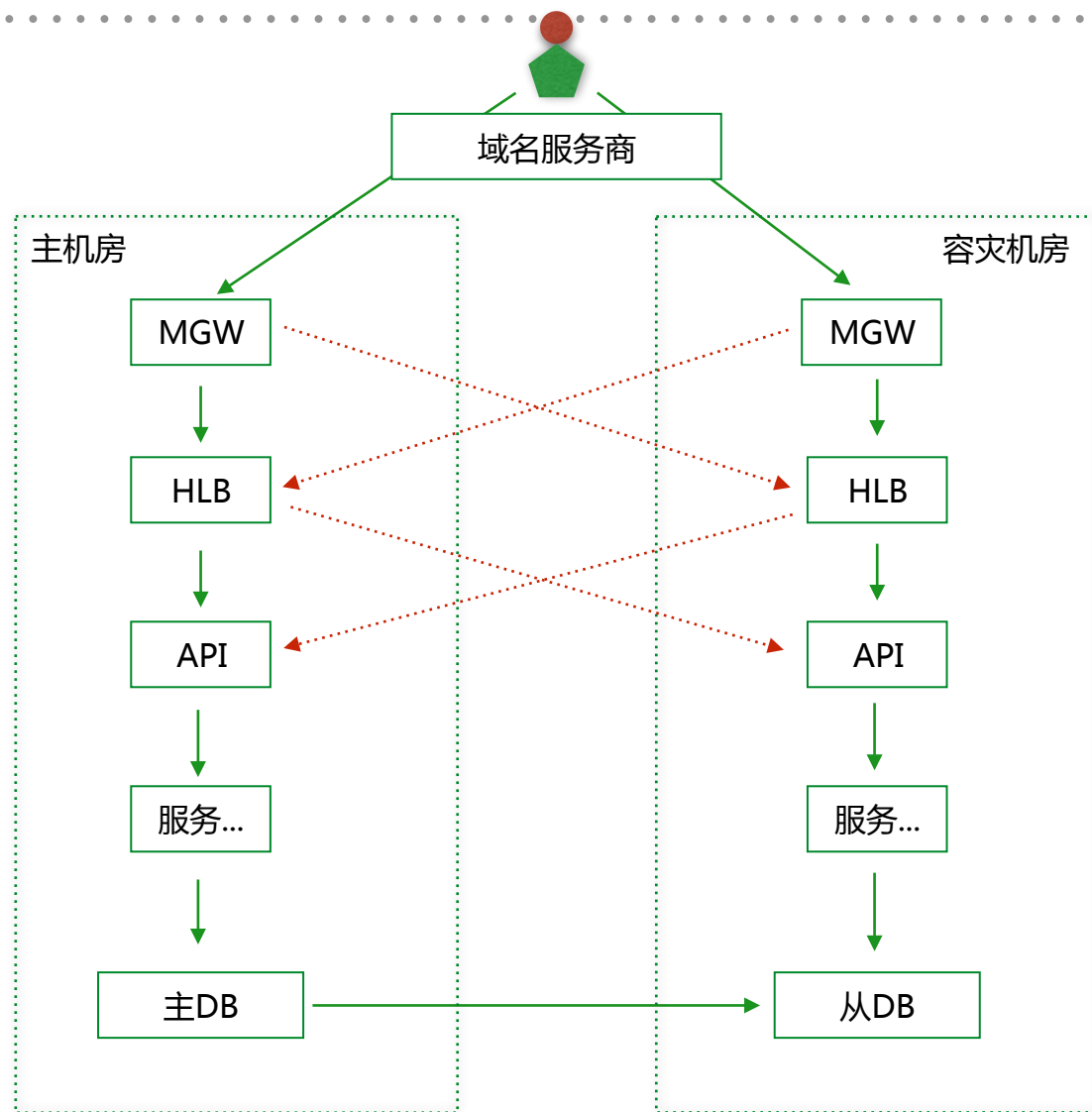
# 架构高可用



- MGW：四层负载均衡网关
  - Http：Http负载均衡及服务治理
  - 服务层：Octo服务治理
  - 数据层：DB、分布式缓存、大数据存储
- ★ 各层在目标容量值下均可按需扩容
- 服务节点
  - 计算节点
  - 存储节点



# 架构高可用：双机房容灾



- 外网故障
- 主机房故障
- 容灾机房故障
- 机房专线故障

高可用：容量规划

容量规划

容量规划举例：DB容量规划

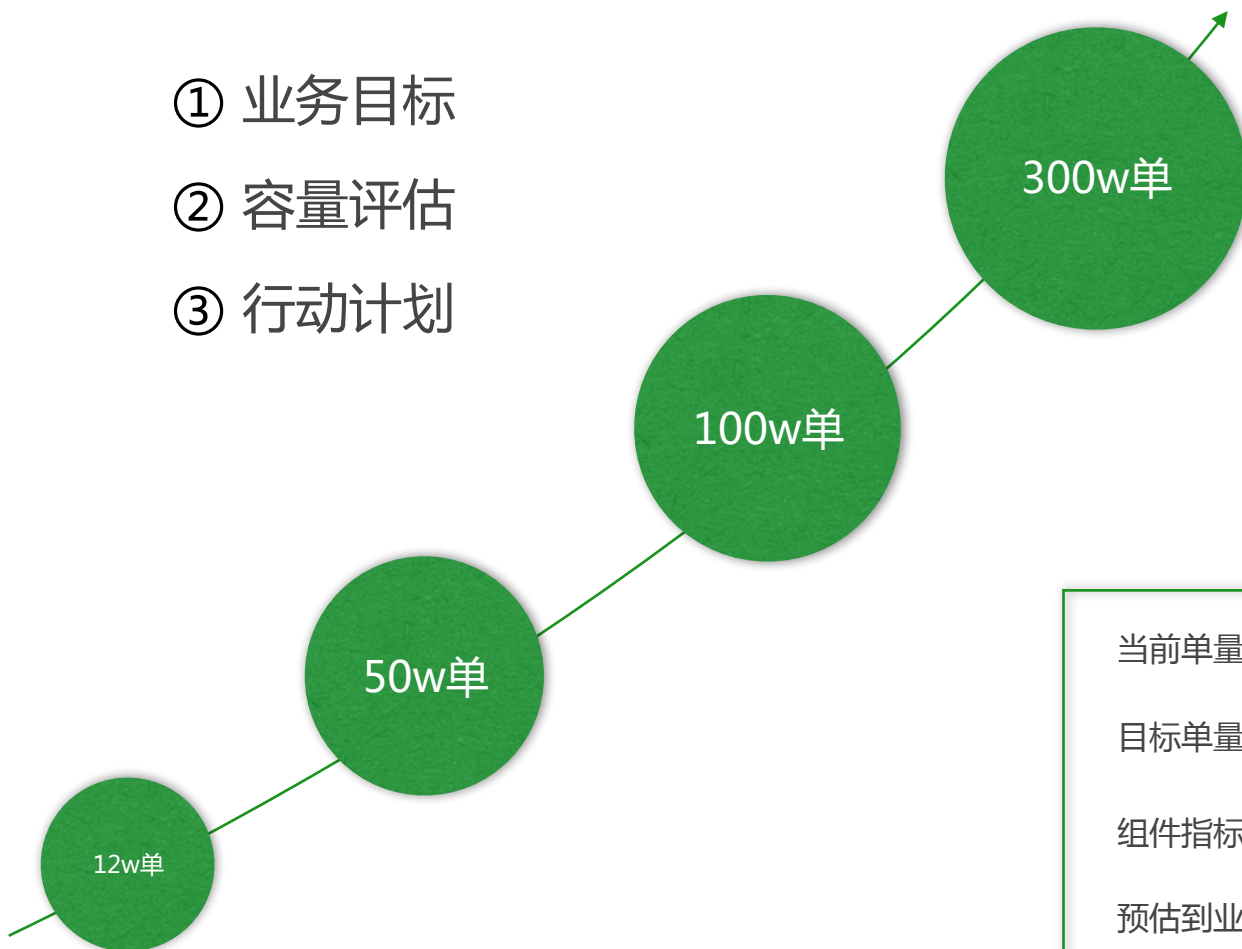
容量预估+压测

# 容量规划

① 业务目标

② 容量评估

③ 行动计划



容量评估：

- 业务维度：单量、活跃骑手数
- 组件容量和业务维度相关性
- 组件指标
- 组件预警值

当前单量：  $m1$  当前活跃骑手数：  $n1$

目标单量：  $m2$  目标活跃骑手数：  $n2$

组件指标值为  $v$ ，其中  $p\%$  和单量相关， $1-p\%$  和骑手数相关

预估到业务目标时的指标值： $v * p\% * m2 / m1 + v * (1 - p\%) * n2 / n1$

# DB容量规划： 50w单

50w单

## 评估时业务现状

业务量： 12w单

DB指标： 主库写qps：690

## 到达目标业务量的DB评估

单量：50w单（ x 4.2 倍）

- 分析sql组成及目标指标值  
2608， 超阈值2000
- 瓶颈：  
骑手坐标在写库占比59.8%

## Action： DB架构升级

- 骑手实时坐标表： 分布式缓存
- 骑手历史坐标表： 拆库、一天一张表

# DB容量规划： 100w单

---

100w单

## 评估时业务现状

---

业务量： 30w单

DB指标：

- 主库写qps：800
- 从库读Qps  
(单从库)：7500

## 到达目标业务量的DB评估

---

100w单 ( x 3.34 )

- 主库写：2672  
瓶颈：track 占比24%
- 单从库读：25050  
瓶颈：人员、组织数据查询占比27%

## Action：DB架构升级

---

- 主库：track拆成单独库 + 归档
- 人员、组织（不易变的数据）：  
分布式 缓存，databus同步
- 监控容量，增加从库

# DB容量规划： 300w单

300w单

## 评估时业务现状

业务量： 70w单

DB指标：

- 主库写qps：1500
- 主库读qps：5500qps

## 到达目标业务量的DB评估

300w单 ( x 4.29 )

- 主库写2793  
瓶颈： 运单表70%、结算表20%
- 主库读2w5左右  
瓶颈： 运单表

## Action： DB架构升级

- 拆库：结算表拆成单独库
- 拆库：运单表拆成单独库
- 主库降低读（从库、缓存）
- 运单表散库散表

# 容量规划

---

- 持续的容量规划，把握住架构升级节奏
  - 太早：影响业务需求迭代速度、运维成本高
  - 太晚：架构升级跟不上业务发展
  - 技术方案长期规划、分阶段逐步实施
- 需求上面小步快跑，架构上面持续重构，小步快跑

# 容量预估 + 压测

---

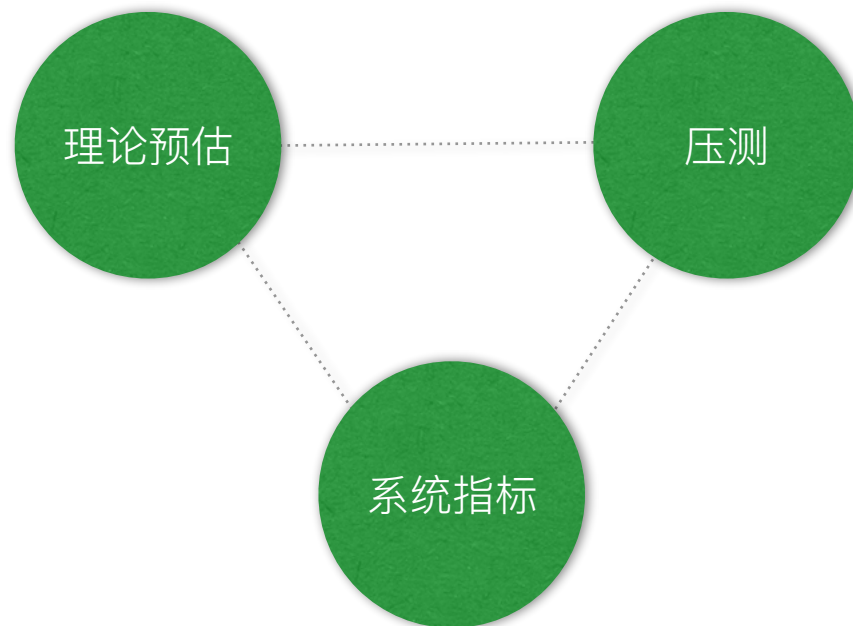
理论预估：

- 定时预估、容量规划
- 新功能上线时，需要对下游服务、组件（缓存、DB）等进行预估

压测

线下压测

线上全链路压测





高可用：降级

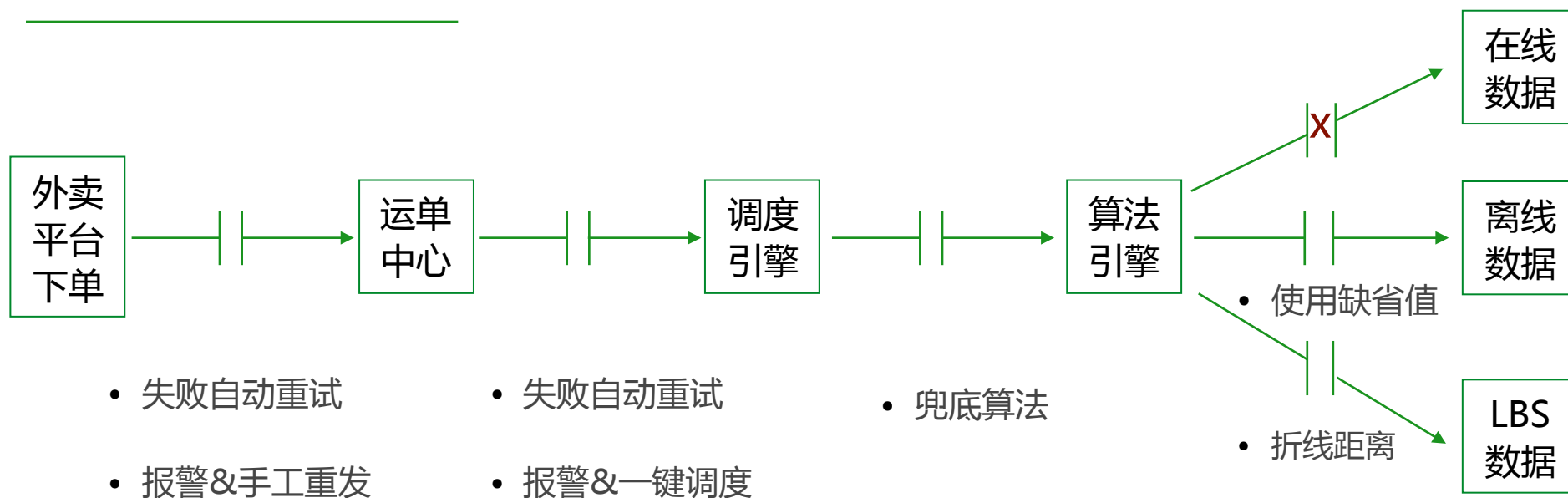
关键路径保障及兜底

二层降级体系

止损

# 降级体系： 关键路径保障及兜底

## 运单调度关键路径兜底



# 降级体系： 二层降级体系

## 二层降级体系

- 按场景预案组合降级
  - 限流组合降级： app防刷 + 接入层限流
- 功能降级
  - 非核心路径
  - 核心路径兜底
  - 异步

2	kv_rider_uuid	一级降级(影响严重)	banma.banma_api	FUNC_KV	检查多终端登录, 读kv取uuid	降级期间, 所有请求不再检查多终端登录, 逻辑同于KV异常	未降级	✎
3	kv_set_lastest_online_time	二级降级(影响可控)	banma.banma_api	FUNC_KV	设置骑手最后在线时间, 写kv	降级期间, 所有请求不再设置骑手最后在线时间, 逻辑同于KV异常	未降级	✎
4	kv_set_lastest_polling_msg	二级降级(影响可控)	banma.banma_api	FUNC_KV	设置骑手的通知列表, 写kv	降级期间, 骑手的polling将取不到该时间段内的通知, 不影响push逻辑, 逻辑同于KV异常	未降级	✎
5	kv_add_full_polling_msg	二级降级(影响可控)	banma.banma_api	FUNC_KV	设置骑手的最新通知, 写kv	降级期间, 骑手的polling将取不到该时间段内的通知, 不影响push逻辑, 逻辑同于KV异常	未降级	✎
6	kv_get_lastest_polling_msg	二级降级(影响可控)	banma.banma_api	FUNC_KV	获取骑手的最新, 读kv	降级期间, 骑手的polling将取不到通知, 不影响push逻辑, 逻辑同于KV异常	未降级	✎
7	kv_get_full_polling_msg	二级降级(影响可控)	banma.banma_api	FUNC_KV	获取骑手的通知列表, 读kv	降级期间, 骑手的polling将取不到通知, 不影响push逻辑, 逻辑同于KV异常	未降级	✎
8	kv_delivery_count_statistics	一级降级(影响严重)	banma.banma_api	FUNC_KV	统计骑手运单量信息, 读写kv	[慎用!!!]降级期间, 骑手运单统计信息完全从db中读取, 每次请求访问数据库	未降级	✎

共100多项功能降级

# 降级体系：止损

## 止损

- 核按钮
- 关店

按业务线关店  
按城市关店  
按区域关店

业务管理 / 批量商家设置

城市

配送区域

区域类型

是否置休

是否下活动

配送费

查询

操作记录

批量设置

活动开启/关闭

商户营业/置休

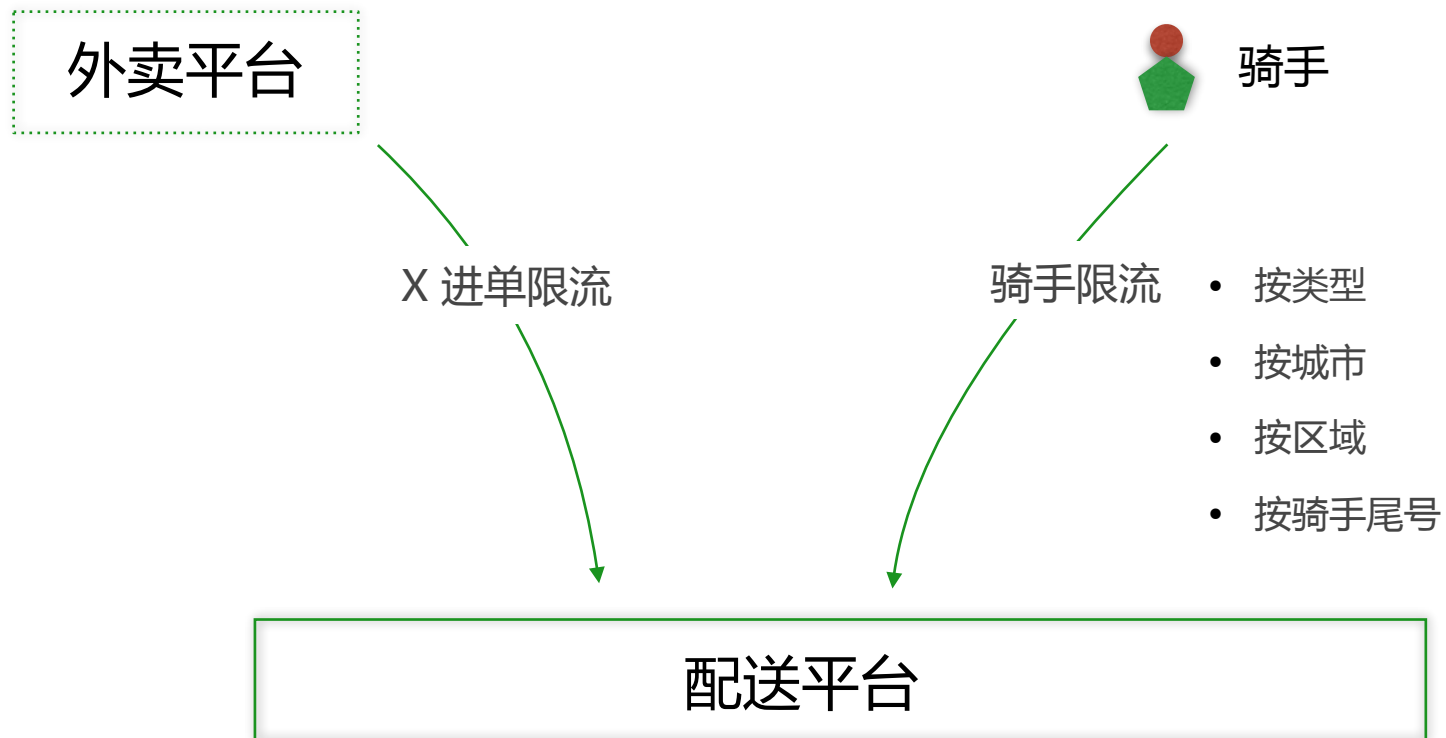
调整配送费

城市	城市ID	区域	区域ID	区域类型	是否置休	是否下活动	配送费调整	操作
青岛	安全隐藏	安全隐藏	安全隐藏	城市代理	否	否	安全隐藏	   
长垣县				城市代理	否	否		   
武汉				城市代理	否	否		   
武汉				城市代理	否	否		   
武汉				城市代理	否	否		   
汕尾				城市代理	否	否		   
洛阳				城市代理	否	否		   
孝感				城市代理	否	否		   
新乡				城市代理	否	否		   
深圳				自营	否	否		   

高可用：限流

# 限流体系

.....



高可用： 监控报警体系

四层监控体系

监控大盘

分级报警

持续完善

# 监控报警体系： 四层监控体系

.....

## 硬件监控

CPU

内存

硬盘

网络

...

## 基础组件监控

DB

消息中间件

DataBus

...

## 服务监控

性能监控

异常监控

JVM监控

...

## 业务监控

关键数据监控

关键链路监控

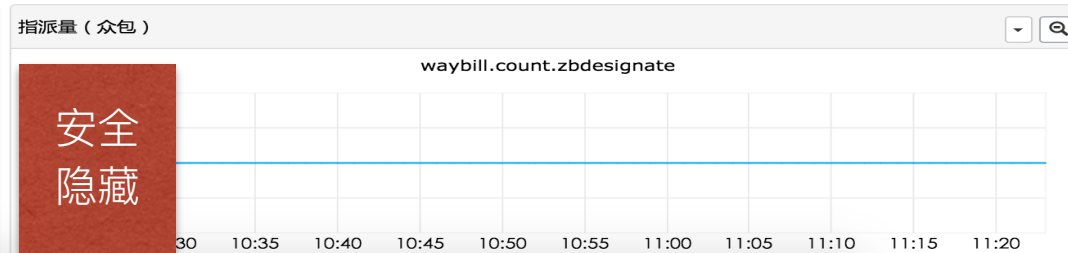
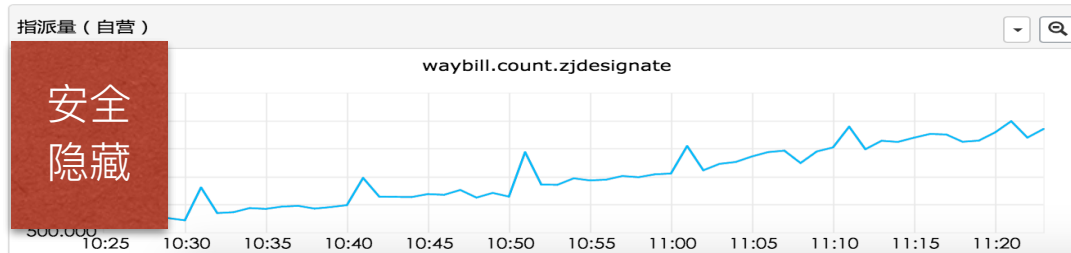
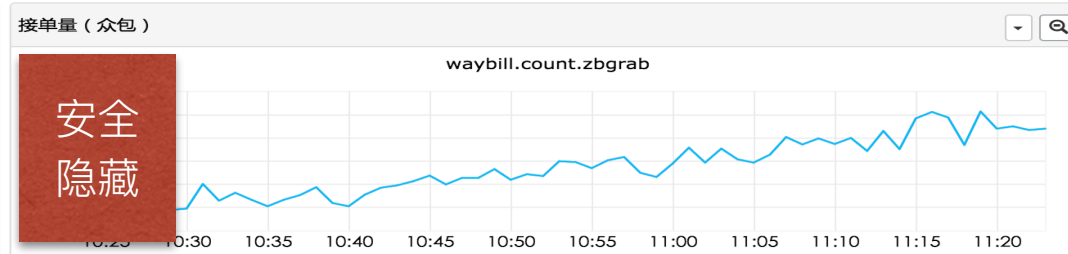
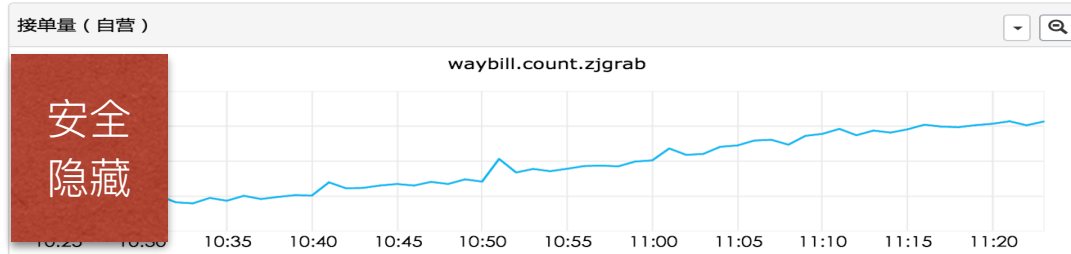
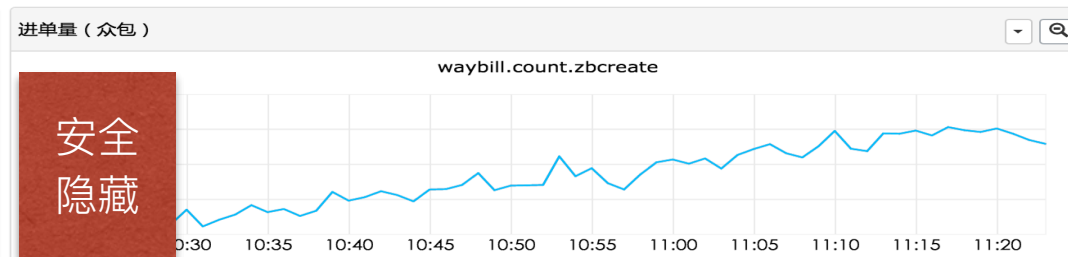
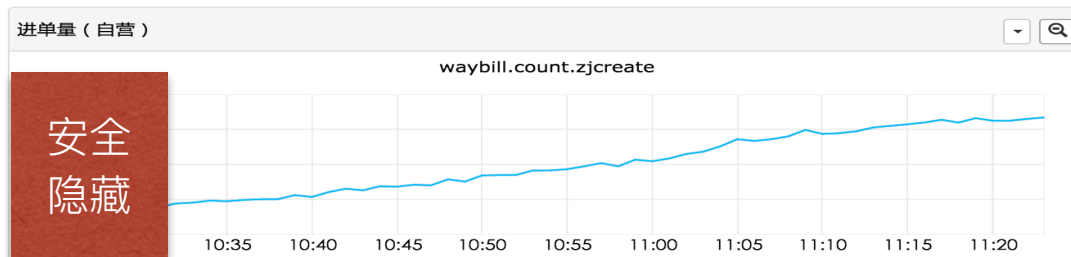
兜底方案监控

...



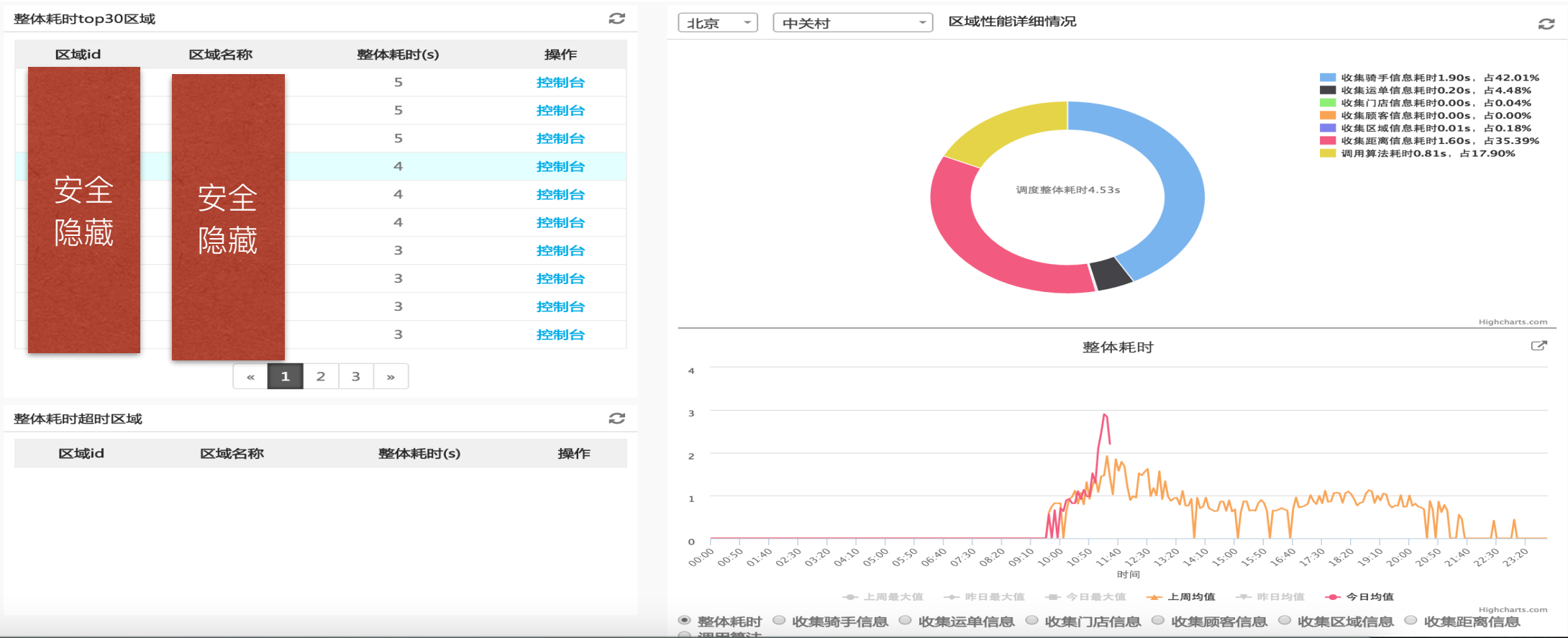
# 监控报警体系：业务监控

- 关键业务数据监控 + 报警（趋势报警、尖刺报警等）



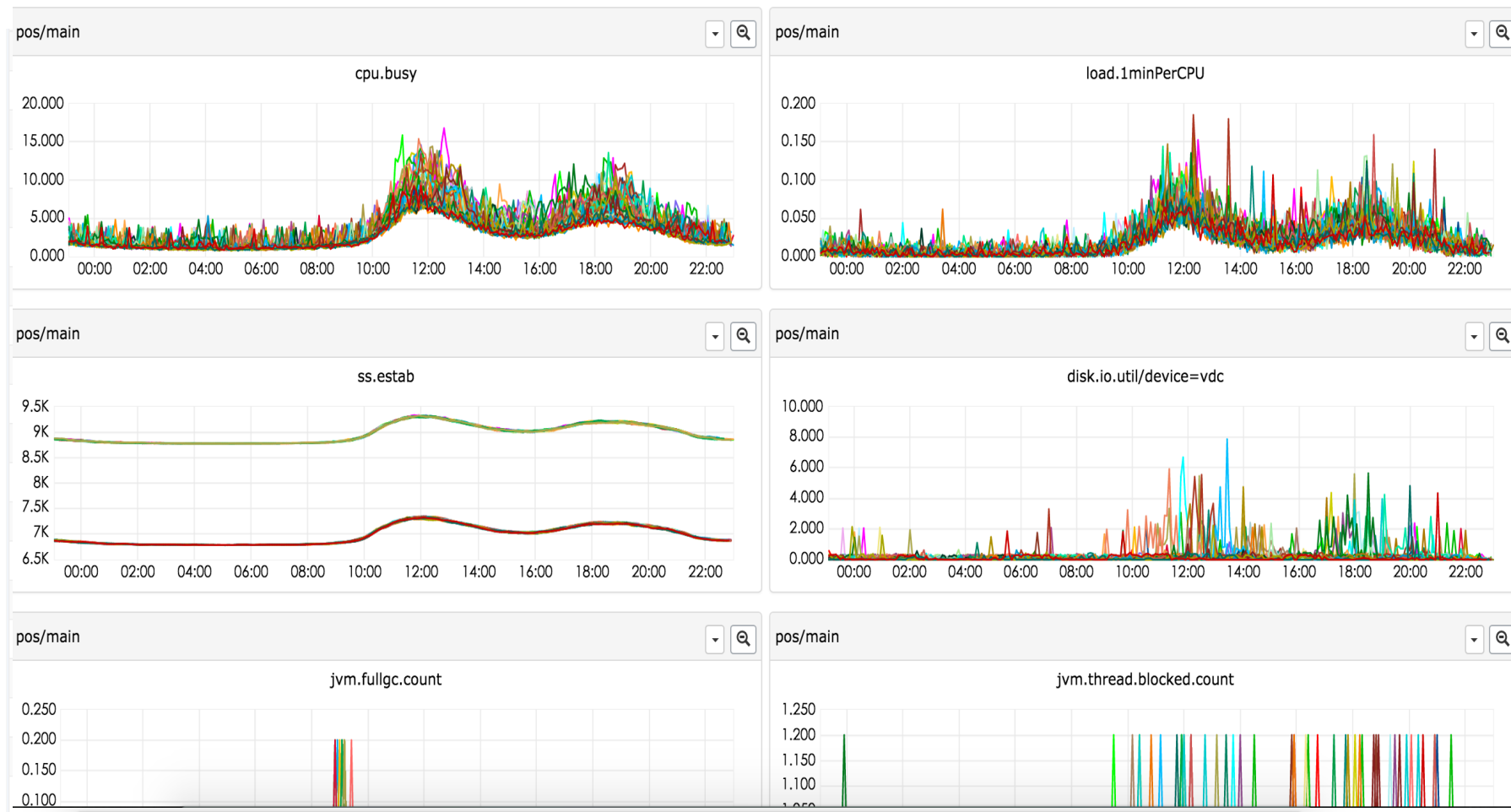
# 监控报警体系：业务监控

- 关键链路\关键节点监控



# 监控报警体系：监控大盘

- 系统监控大盘
- 服务监控大盘



# 监控报警体系：持续完善

---

- 定期Review
- case驱动

高可用：流程

# 流程保障

.....

## 事前预防

---

- 上线流程
- 数据库变更
- 洗数据
- 容量预估
- 扩容流程
- 代码Review
- 灰度发布流程

## 事中处理

---

- 线上故障处理流程

## 事后回顾

---

- CaseStudy

高压线

谢谢大家

